BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 2: e25876
doi: 10.3897/biss.2.25876

OPEN ACCESS

Conference Abstract

Workflow and Current Achievements of BIOfid, an Information Service Mobilizing Biodiversity Data from Literature Sources

Christine Driller[‡], Markus Koch[§,|], Marco Schmidt[§,¶], Claus Weiland[§], Thomas Hörnschemeyer[#], Thomas Hickler[§], Giuseppe Abrami[¤], Sajawel Ahmed[¤], Rüdiger Gleim[¤], Wahed Hemati[¤], Tolga Uslu[¤], Alexander Mehler[¤], Adrian Pachzelt[«], Jashar Rexhepi[«], Thomas Risse[«], Janina Schuster[«], Gerwin Kasperek[«], Angela Hausinger[«]

‡ Senckenberg Gesellschaft für Naturforschung, Frankfurt am Main, Germany
§ Senckenberg Biodiversity and Climate Research Centre, Frankfurt am Main, Germany
| Institute of Evolutionary Biology and Ecology, University of Bonn, Bonn, Germany
¶ Palmengarten der Stadt Frankfurt, Frankfurt am Main, Germany
# Senckenberg Research Institute and Natural History Museum Frankfurt, Frankfurt am Main, Germany
¤ Faculty of Computer Science and Mathematics, Goethe-University Frankfurt, Frankfurt am Main, Germany
« University Library Johann Christian Senckenberg, Goethe-University Frankfurt, Frankfurt am Main, Germany

Corresponding author: Christine Driller (christine.driller@senckenberg.de)

Received: 16 Apr 2018 | Published: 16 Apr 2018

Citation: Driller C, Koch M, Schmidt M, Weiland C, Hörnschemeyer T, Hickler T, Abrami G, Ahmed S, Gleim R, Hemati W, Uslu T, Mehler A, Pachzelt A, Rexhepi J, Risse T, Schuster J, Kasperek G, Hausinger A (2018) Workflow and Current Achievements of BIOfid, an Information Service Mobilizing Biodiversity Data from Literature Sources. Biodiversity Information Science and Standards 2: e25876.
https://doi.org/10.3897/biss.2.25876

Abstract

BIOfid is a specialized information service currently being developed to mobilize biodiversity data dormant in printed historical and modern literature and to offer a platform for open access journals on the science of biodiversity. Our team of librarians, computer scientists and biologists produce high-quality text digitizations, develop new text-mining tools and generate detailed ontologies enabling semantic text analysis and semantic search by means of user-specific queries. In a pilot project we focus on German publications on the distribution and ecology of vascular plants, birds, moths and butterflies extending back to the Linnaeus period about 250 years ago. The three organism groups

© Driller C et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

have been selected according to current demands of the relevant research community in Germany. The text corpus defined for this purpose comprises over 400 volumes with more than 100,000 pages to be digitized and will be complemented by journals from other digitization projects, copyright-free and project-related literature. With *TextImager* (Natural Language Processing & Text Visualization) and *TextAnnotator* (Discourse Semantic Annotation) we have already extended and launched tools that focus on the text-analytical section of our project. Furthermore, taxonomic and anatomical ontologies elaborated by us for the taxa prioritized by the project's target group - German institutions and scientists active in biodiversity research - are constantly improved and expanded to maximize scientific data output. Our poster describes the general workflow of our project ranging from literature acquisition via software development, to data availability on the BIOfid web portal (http://biofid.de/), and the implementation into existing platforms which serve to promote global accessibility of biodiversity data.

Keywords

Biodiversity, Specialized Information Service, Ontologies, Text mining

Presenting author

Christine Driller

Presented at

SPNHC + TDWG Conference 2018

Funding program

Acquisition and Provision (Scientific Library Services and Information Systems DFG-LIS)

Grant title

Specialized Information Service for Biodiversity Research (DFG-Projectnumber 326061700)